

CLAIMS

- 1 1. A method for initiating a peer-to-peer communication session, the method comprising the steps of:
2 attempting a first remote direct memory access (RDMA) read operation directed
3 to a cluster partner;
4 performing, in response to a successful first RDMA read operation, a first RDMA
5 write operation to the cluster partner;
6 performing, in response to a successful RDMA write operation, a second RDMA
7 read operation directed to the cluster partner; and
8 performing, in response to a successful second RDMA read operation, a second
9 RDMA write operation to the cluster partner.
10
- 1 2. The method of claim 1 wherein the step of attempting a first RDMA read operation further comprises the step of issuing a RDMA read operation to the cluster partner requesting a pre-set memory address location that is associated with a status variable on the cluster partner.
- 1 3. The method of claim 1 further comprising the steps of:
2 exchanging a set of peer connection information;
3 passing a set of client information to the cluster partner;
4 creating a set of appropriate communication ports;
5 alerting the cluster partner of a ready status; and
6 alerting a set of clients that the cluster partner is in a ready state.
- 1 4. The method of claim 3 wherein the set of peer connection information comprises
2 a version number.
- 1 5. The method of claim 1 wherein the step of passing a set of client information to the cluster partner further comprises the steps of:
2 collecting, from a set of clients, the set of client information; and
3

- 4 transferring the collected set of client information to the cluster partner.
- 1 6. The method of claim 5 wherein the client information comprises a number of
2 communication ports required.
- 1 7. The method of claim 5 wherein the set of client information further comprises an
2 amount of memory requested by a particular client.
- 1 8. The method of claim 1 wherein the cluster partner is a storage system.
- 1 9. The method of claim 1 wherein the cluster partner is an application server.
- 1 10. A storage operating system, executing on a storage system, the storage operating
2 system comprising:
3 a cluster connection manager adapted to initiate a peer to peer communication
4 session with a cluster partner upon initialization of the storage operating system.
- 1 11. The storage operating system of claim 10 wherein the cluster connection manager
2 further comprises:
3 means for performing a remote first direct memory access (RDMA) read opera-
4 tion directed to a cluster partner;
5 means for performing, in response to a successful first RDMA read operation, a
6 first RDMA write operation to the cluster partner;
7 means for performing, in response to a successful first RDMA write operation, a
8 second RDMA read operation directed to the cluster partner; and
9 means for performing, in response to a successful second RDMA read operation,
10 a second RDMA write operation to the cluster partner.
- 1 12. The storage operating system of claim 11 wherein the cluster connection manager
2 further comprises:
3 means for exchanging a set of peer connection information;

- 4 means for passing a set of client information to the cluster partner;
- 5 means for creating a set of appropriate communication ports;
- 6 means for alerting the cluster partner of a ready status; and
- 7 means for alerting a set of clients that the cluster partner is in a ready state.

1 13. A method for initiating a peer-to-peer communication session, the method com-
2 prising the steps of:

- 3 performing a first remote direct memory access read operation directed to a clus-
4 ter partner; and
- 5 performing, in response to a successful first remote direct memory access read
6 operation, a first remote direct memory access write operation to the cluster partner.

1 14. The method of claim 13 wherein the first remote direct memory access read op-
2 eration is performed over a Virtual Interface connection having a pre-determined and pre-
3 assigned Virtual Interface Number and a pre-determined Fibre Channel ID.

1 15. A method for initiating a peer-to-peer communication session, the method com-
2 prising the steps of:

- 3 (a) attempting a first remote direct memory access read operation directed to a
4 predefined hardware address and a predefined port number; and
- 5 (b) performing, in response to a successful step (a), a first remote direct memory
6 access write operation directed to the predefined hardware address and the predefined
7 port number..

1 16. The method of claim 16 further comprising the step of:

- 2 (c) performing; in response to a successful step (b), a second remote direct mem-
3 ory access read operation directed to the predefined hardware address and the predefined
4 port number.

1 17. The method of claim 15 wherein the predefined hardware address comprises a
2 fibre channel identifier.

1 18. The method of claim 15 wherein the predefined port number comprises a virtual
2 interface.

1 19. The method of claim 15 wherein the first remote direct memory access is deliv-
2 ered to a predefined memory address storing booting status information.

1 20. A system configured to establish reliable peer-to-peer communication among storage
2 systems of a clustered environment, the system comprising:
3 a peer process executing on each storage system partner; and
4 a cluster connection manager executing on each storage system partner, the clus-
5 ter connection manager establishing a reliable peer-to-peer connection between each peer
6 process by connecting to a predetermined port number using a predetermined network
7 address.

1 21. The system of claim 20 wherein the reliable peer-to-peer connection is established
2 without requiring a storage operating system executing on each storage system partner to
3 be fully functioning.

1 22. The system of claim 20 wherein the peer-to-peer connection is a virtual interface
2 connection.

1 23. The system of claim 20 wherein the peer process is a cluster connection client that
2 requests services from the cluster connection manager.